Foundations of an Interactive Activation Model of Eye Movement Control in Reading

Ronan G. Reilly
Department of Computer Science
National University of Ireland, Maynooth

Ralph Radach
Department of Psychology
Technical University of Aachen

Abstract

This chapter describes an interactive activation model of eye movement control in reading, which we refer to as "Glenmore", that can account within one mechanism for preview and spillover effects, and for regressions, progressions, and refixations. The model decouples the decision about when to move the eyes from the word recognition process. The time course of activity in a "fixate centre" determines the triggering of a saccade. The other main feature of the model is the use of a saliency map that acts as an arena for the interplay of bottom-up visual features of the text, and top-down lexical features. These factors combine to create a pattern of activation that selects one word as the saccade target. Even within the relatively simple framework proposed here, a coherent account has been provided for a range of eye movement control phenomena that have hitherto proved problematic to reconcile.

Theoretical background and an outline of the model

Current models of eye movement control in reading.

In recent years, research on eye movements in reading has made substantial progress. A key new development in the field is the emergence of computational models of eye movement control during reading (see Kennedy, Radach, Heller & Pynte, 2000 and Reichle, Rayner & Pollatsek, in press, for detailed discussions). The modeling principles and the algorithms that these models implement reflect the theoretical viewpoints of their authors. For example, in the influential E-Z Reader model (Reichle, Pollatsek, Fisher & Rayner, 1998) sequential lexical processing is suggested to be the obligatory trigger for the generation of all eye movements made in normal reading. In contrast, Reilly & O'Regan (1998), following the theoretical framework developed by O'Regan (1990), attempted to demonstrate that a good account for the positioning of fixations in reading can be achieved by using a set of rather dumb oculomotor heuristics. We believe that both of these positions have their merits and can account for important aspects of eye behaviour during reading. On the other hand, both approaches have also serious limitations.

It is quite clear that a pure visuomotor account as proposed by Reilly and O'Regan (1998) is not sufficient to explain many phenomena that are apparent in human reading behaviour. This has been acknowledged already by O'Regan, Vitu, Radach and Kerr (1994) who suggested that visuomotor and cognitive theories of eye movement control in reading will need to be combined: "The resulting intermediate theory contains both an underlying scanning strategy that can manifest itself even in the absence of linguistic material to process; it contains a modulator influence of linguistic processing… Future work will show how exactly these components must be combined." (p. 345).

Thus, the question of interest is not whether eye movements are determined by visuomotor factors or linguistic processing, but to what degree these two types of factors are involved and how they interact. Taking the likelihood of fixating a word as an example, Brysbeart and Vitu (1998) have shown in an elegant meta-analysis that most of the variance in "word skipping" is accounted for by word length.[1] At the same time there are significant influences from two cognitive factors, word frequency and contextual predictability. A

[1] It could be argued that word length carries linguistic information, for example, in terms of reducing the number of word candidates activated in parafoveal preprocessing. However, as Inhoff, Radach, Eiter and Juhasz (in press) have recently shown, this appears not to be the case.

similar example is the analysis of refixations by McConkie, Kerr, Reddix, Zola, and Jacobs (1989), who first showed that the likelihood of immediately refixating a word based on initial fixation position can be expressed as a u-shaped (quadratic) function. Importantly, only the vertical shift parameter of the refixation function varies with word frequency, while the slope of the function is determined by a visual factor, the eccentricity of the initial fixation position relative to the word centre. Analyses of this type appear to suggest that eye viewing behaviour is co-determined by low-level visual and higher-level cognitive factors, possibly in terms of low-level default routines that are affected substantially by cognitive modulation.

The critique of models that are based on sequential lexical processing and discrete shifts of attention has taken different directions. From a computational point of view, Engbert and Kliegl (2001) have shown that about the same fit of empirical reading data can be achieved using a model that includes simpler principles of operation. Their model differs from E-Z Reader in at least two important respects: the two phases of the word recognition process (familiarity check and lexical access) are replaced with an all-or-none process, and sometimes "autonomous" saccades are triggered independent of lexical processing.

A second line of criticism is based on empirical work suggesting that word processing during reading may be spatially distributed rather than confined to one word at a time as suggested by sequential attention shift models. These observations usually take the form of demonstrating that characteristics of a prefoveal word can influence the duration of a fixation on the currently fixated foveal word (e.g., Kennedy, 1998; Kennedy *et al.,* in press). Inhoff, Radach and Greenberg (2000) have shown that information is regularly being acquired from positions left of the currently fixated word. Moreover, a recent experiment by Inhoff, Lemmer, Eiter and Radach (2001) suggests that information from a parafoveal word can be acquired very early during a reading fixation. Although these results may not yet be conclusive, several sources of evidence point to the possibility that linguistic processing during reading can operate on (at least) two words in parallel.

This view is corroborated by another line of criticism dealing with the time line of processing events. As discussed by Deubel, O'Regan, and Radach (2000) and Radach, Inhoff and Heller (2002), basic oculomotor research using the double step paradigm indicates that the (re-)programming of a saccade must be initiated at least 70 ms to 90 ms before the end of the current fixation. This is similar to the suggestion by McConkie, Underwood, Zola and Wolverton (1985) of an interval of 80-100 ms before saccade onset as the deadline for stimulus influences during a fixation. Sereno, Rayner and Posner (1998) report ERP results in a word recognition paradigm indicating that, starting at 132 ms after stimulus onset, word frequency differences can be observed. Since the lexical familiarity stage in the E-Z Reader Model is sensitive to word frequency effects, this phase of word processing cannot take place much earlier than the 132 ms estimate. Taken together, the conclusion is that the effective time window for direct (immediate) linguistic influences during a fixation of 250 or 275 ms is very limited. Looking at the scenario that sequential attention models suggest to account for “word skipping” (see e.g., Pollatsek, Reichle, & Rayner, this volume), it is not clear how operations like lexical access on the origin word, a shift of attention to the skipped word and a familiarity check on the skipped word could all be completed in sequential way within this brief time interval.

One way to accommodate the problems discussed above is to relax assumptions concerning the allocation of attention, allowing more then one word to be lexically processed concurrently. Following LaBerge and Brown (1989), Inhoff, Radach, Starr and Greenberg (2000) suggested replacing the idea of a sequentially moving attentional spotlight by an attentional gradient around the point of fixation that can include several words. A similar route has been taken by Engbert, Longtien and Kliegl (2002) who developed a computational model of eye movement control in reading that incorporates the notion of spatially distributed lexical processing. Other core principles of their model are an autonomous timing of saccades that tend to be generated at a preferred rate and the inhibition of saccade initiation by foveal lexical processing. On the other hand, Engbert, Longtien and Kliegl (2002) have maintained some key features of the original Easy Reader model, for example, a division of lexical processing into an early versus late stage, and a distinction between a labile and a non-labile phase of saccade programming (see Kliegl & Engbert, this volume, for a discussion).

<Insert Figure 1 about here>

An outline of our theory and model

The model that will be described in this chapter represents a more radical departure from the sequential attention shift conception. While our theory has some apparent similarities with the SWIFT model and the competition-inhibition model by Yang and McConkie (2001), there are also important differences.

Our theory and its first computational implementation, the "Glenmore" model[1], are closely related to the general theory of saccade generation developed by Findlay and Walker (1999). They proposed that saccade target selection is accomplished via parallel processing and competitive inhibition within a two-dimensional salience map. The triggering of a saccade is controlled by a fixate centre that is sensitive to input from several routes of cognitive processing. Once a saccade is triggered, it will go to the target that has emerged as a winner in the saliency competition. Radach (1999) has proposed that the saliency map mechanism could provide an elegant basis for a theory of saccade generation in reading, avoiding many complexities inherent to the class of sequential attention models. In particular, we deviate from the view that the word-by-word sequence within a line of text provides a chain of default saccade targets and that the observable eye movement behaviour is generated via frequent cancellation and reprogramming of default saccades.

In most cognitive activities that require systematic visual scanning, every saccade is directed to a specific target object. In the case of reading this selection takes the form of deciding which of the words within the current "perceptual span" should be the target for the next saccade. In the vast majority of cases, this will include the word currently fixated, the preceding word, and the two or three following words. The most important low-level sources of influence on this decision are the length and the eccentricity of words located around the current point of fixation (Kerr, 1992; McConkie, Kerr & Dyre, 1994). Following Findlay and Walker, we assume that potential target words are represented (perhaps as low spatial frequency objects) on a salience map and, depending on the particular visual configuration, their salience values form a preference list of potential targets.

We further assume that at the beginning of each fixation low level visual information, coded as a saliency vector, is available that allows for the triggering of a saccade without any

[1] The name Glenmore originates from the location of a remote cottage in the south west coast of Ireland, where the foundations of our modeling work were laid during a one-week retreat there.

cognitive influence.[1] Over time, the saliency values representing potential targets will change in response to information about ongoing linguistic processing. This simple principle can be illustrated using the following example: Suppose that a reader is fixated on a letter in the right half of a word of medium length, the next word (N+1) is short and the word N+2 again of medium size. Given this visual configuration, it is likely that word N+2 will have the highest initial salience value and will be the target for the next interword saccade. If however, word N+1 turns out to be difficult to process, its saliency value may rise quickly and it will become a more attractive target. As will be discussed later, in this scenario there is competition between words for limited processing resources, opening a route to explain fovea-on-parafovea and parafovea-on-fovea effects. The above example shows also that within a spatial saliency framework the notion of "word skipping" becomes meaningless, as there are is no default saccade program to N+1 that needs to be cancelled and reprogrammed (see Brysbeart & Vitu, 1998, for a similar idea).

In addition to the saliency map representation, the Glenmore model includes a visual input module, a word processing module, a fixate centre and a saccade generator, producing the actual saccadic movement (see Figure 1). A visual input vector codes the visual configuration around the current fixation position. The computation of initial saliency values is based on a letter processing function (McConkie & Zola, 1987) and effectively accounts for effects of word length and eccentricity. Visual information is transferred to the saliency map representation and to a linguistic processing module that implements processing on the letter and word level within an interactive activation (IA) framework (Grainger & Jacobs, 1998). From the linguistic processing module information about ongoing processing is transferred in two directions. The vector of letter unit activation is transmitted to the saliency map, where it is used to update continuously the saliency values of potential saccade targets.

---

[1]This idea is related to the notion of a visual scanning routine as proposed by Levy-Schoen (1981). She suggested that "The routine does not determine absolute saccade length but rather criteria according to which saccade length will be programmed so that the eyes move in a way relevant to the task" (p. 301). It is assumed that readers learn scanning patterns individually as a way to navigate through configurations of low spatial frequency word objects. During this process, a reader will develop routines that, on average, provide optimal information acquisition for letter and word processing. Our implementation represents a first approximation to explicitly modeling this mechanism.

At the same time, feedback on the general level of excitation in the word processing network is sent to the fixate centre. The triggering of a saccade is based on activity in a fixate module that operates in relation to the dynamics of spatial saliency. Over the course of a fixation, activity in the fixate centre will tend to fall, a process that has a random component (similar to autonomous saccade triggering in SWIFT or the random waiting time component in Yang & McConkie's model) and a non-spatial processing component. The saccade will be executed by the saccade generator module after a latency period and will always be directed to the target with maximum saliency. The saccade generator implements the front end behaviour of the eyes as described by McConkie, Kerr, Reddix and Zola (1988) and implemented in Reilly and O'Regan (1998) and Reichle, Rayner, and Pollatsek (1999).

We believe that one important feature of our spatial saliency theory of eye movement control in reading is its neuroscientific plausibility. The general architecture is in harmony with neurobiological constraints and information processing principles suggested by oculomotor research (see e.g., Wurtz, 1996; Carpenter, 2000; Munoz, this volume). An important feature of the Glenmore model is that it operates on the level of individual letters *and* words both in terms of visual and linguistic processing and eye movement control. The inclusion of a realistic interactive activation network of letter and word processing is motivated by the fact that IA models have proven especially useful for capturing the time course of parallel activation and competitive inhibition between processing units on a hierarchy of levels. We see the letter and word processing part of the model as a step towards the necessary integration of modeling efforts in the neighbouring domains of word recognition and continuous reading (Grainger, 2000; Jacobs 2000).

It is also worth pointing out that the spatial saliency theory of eye movement control in reading does not refer to the concept of "visual attention" in any way. We agree with Findlay and Walker (1999) who noted that not much is gained by assuming that "attention" is disengaged, moved and re-allocated as a function of saliency and that it is these processes which, in turn, trigger saccade programming. From an epistemic point of view it appears that proposing that "attention" moves from word to word is only shifting the problem of explanation to another level: if attention is supposed to be an entity that moves somewhat independently from eye movements, this *movement* will not only need to be triggered. It will also need to be programmed, it will need to have a targeting mechanism, a latency and a duration, requiring a machinery of "attention generation" to co-exist concurrently with that of saccade generation.

Before the details and dynamics of the model are described in more detail, we will discuss some general issues related to computational modeling in the area of eye movement control in reading.

Computational Modelling

As mentioned in the introduction, there has been a significant increase in the use of computational modelling techniques to explore in a more rigorous fashion various theories of eye movement control in reading. We believe that this has helped to clarify some important theoretical issues and to eliminate some proposals that proved less viable when put under computational scrutiny. The traditional version of the Morrison (1984) model comes to mind here. Its assumptions regarding the interplay of the time course of saccadic programming and lexical identification proved unsupportable when implemented computationally (Reilly & O'Regan, 1998), which may have contributed to motivating the revised version of Morrison's model embodied in EZ-Reader (Reichle et al., 1998).

Computational modelling serves to clarify theories, but cannot of itself resolve conflicts between them. Guidelines or meta-principles of computational modelling are required that will allow, among other things, the ready comparison of one model with another. A significant weakness of the current state of the field is in the area of model comparison, and the lack of an agreed methodology for doing so. Current computational models are usually compared on the basis of how well they fit a given set of data and how parsimoniously they do so. Authors usually describe the success of their model in terms of reproducing many of the global features of eye movements in reading – fixation and gaze durations, word skipping, refixations etc., all as a function of word processing difficulty. However, when it comes to comparing the performance of models of differing complexity, with, for example fewer free parameters, we see the limitations of this approach to model comparison. There is a need here for a more in-depth comparison than the admittedly important one of comparing how well the models fit the data. At the very least, one should describe how many parameters are used in the model, how they were estimated, and how the numbers compare to the competing model. While an appeal to parsimony may not be entirely appropriate for models that ultimately rest upon a biological foundation (one suspects that evolution is not always transparently parsimonious), some comparison of numbers of parameters and their motivation would be helpful. Even here, the issue of what parameters are relevant to a comparison is moot. For example, if we compare E-Z Reader (Reichle *et al.*,

1998) to a connectionist model (e.g. Reilly, 1993a,b), do we count the modifiable weights of a neural network as free parameters? While providing answers to these questions is beyond the scope of the present discussion, these are issues that need to be tackled, particularly with the proliferation of computational models in the area.

The Glenmore Model

Background

The brain is a computational device best formalised by the *differential* rather than the *propositional* calculus. Dynamical systems theory is increasingly exploited as a means of understanding brain function both at a neural and cognitive level. In cognitive science, a field which has been traditionally dominated by a paradigm in which cognition is taken to be the manipulation of internal symbols, limitations in the symbol-based approach have become increasingly apparent. Researchers throughout cognitive science have been casting around for alternative theoretical frame-works. One of the most productive of these is the dynamical systems framework, according to which cognitive processes are behavioural patterns of non-linear dynamical systems and are best studied using the mathematics of dynamical modelling and dynamical systems theory (Port & Van Gelder, 1995; Kelso, 1995).

A dynamical systems approach has the potential to cast new light on an old issue in the eye movement control literature: What is the role of linguistic factors in the moment-to-moment control of eye movements in reading? Rather than facilitating all-or-none style explanations, a dynamical systems account can potentially accommodate explanations that argue for the interplay, over time, of linguistic and oculomotor factors. Such accounts might have oculomotor facts playing a dominant role early in the processing of the visual input, with lexical and linguistic factors entering the picture further into the processing time line.

One of the goals of Glenmore is to explore this class of dynamical model, one that allows the interplay of factors from multiple levels of representation. The most appropriate class of modeling frameworks for this approach would seem to be connectionist models, and specifically interactive activation (IA) models (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). A typical IA model comprises a set of interconnected neuron like units. Activity is transmitted through the network over weighted connections. The units comprising the network implement a transfer function that combines the unit's inputs and generates an output based on these combined inputs. The nature of the transfer function varies from model to model, but the general design philosophy is to keep the operation it implements relatively simple. So, a typical transfer function might take the weighted sum of inputs and perform some sort of normalising operation using, for example, the sigmoid function. The network "computes" by circulating activation (i.e., real valued numbers) throughout the network until some stopping criterion has been reached. This might be the achievement of a threshold, or the stabilising of levels of activity.

Architecture

The connectionist architecture of the Glenmore model is relatively simple, comprising input units, letter units, saliency units, and word units (see Figure 1). In addition, there is a "fixate centre" unit that controls the decision when to execute a new saccade. This decision is based on the general level of activity in the letter units, once activation in the fixate centre falls below a given threshold, a saccade is targeted to the most salient word "blob" on the saliency map.

As already mentioned, each class of unit has an associated transfer function that determines what kind of output it generates from its inputs, and how this changes over time. The model uses two transfer functions: Gaussian and sigmoid. The Gaussian transfer function allows the unit to accumulate input, such that the output of the unit rises and decays over time. The precise rate of change is determined by the shape of the distribution, which in turn is determined by the two parameters of the Gaussian (mean and standard deviation). These parameters are fixed for the model described here (m=50, sd = 0.3m), such that the output of the function is 1.0 for m=50. The sigmoid transfer function operates in a similar way to the Gaussian, except that its output does not decay over time.

The only parameters free to vary are the weights connecting the units, and even here all weights of the same type (e.g., letter-to-word weights) are given the same value. These variable model parameters are selected by a parameter fitting process based on the Alopex learning algorithm (Unnikrishnan & Venugopal, 1994). More details of the parameter fitting will be given in the section below.

<Insert Figure 2 about here>

A more detailed picture of the model is given in Figure 2. The inputs from the fixation, in the form of a 30 spaces wide "perceptual span" vector of 1s and 0s (indicating the presence or absence of a letter) are fed forward to the letter units. The letter units form the nexus of the processing network, connecting to the word and saliency units. The word units then act as a source of top-down support for the letter units, augmenting the letter activations. The saliency units preserve the spatial representation (or map) of the input, and are the representational structure used in the selection of a saccade target. A saccade is triggered when activation in the fixate centre (FC) unit falls below an adjustable threshold. FC activity is a function of the global level of activity of the letter units. The threshold of the fixate centre can be adjusted on the basis of strategic factors such as reading task and difficulty of the material. Variations in this threshold thus permit the early or late triggering of saccades.

INPUT UNITS

The architecture assumes a visual field of 30 character spaces, with the fovea at position 11. An asymmetric perceptual span is implemented by the probability density function of the gamma distribution centred on the fovea (see Equation 4). The function is used to scale the inputs, where the presence of a character is initially given a value of 1.0, and then scaled down as a function of distance from the fovea.

$$i(x) = \Gamma(x, 3.5, 4.0) \qquad (1)$$

LETTER UNITS

The letter units receive bottom-up activation from the input units, and top-down activation from the word units. The letter unit transfer function is the probability density of a Gaussian distribution.

$$g(x; m, sd) = \frac{1}{\sqrt{2\pi}.sd} e^{-\frac{(x-m)^2}{2.sd^2}} \qquad (2)$$

Where x is the accumulated net input to the unit, m= 50, and $sd = 0.3m$. Note that a given x_i at time $t+1$ is the accumulated weighted sum of the inputs to the unit calculated as:

$$x_{i,t+1} = x_{i,t} + \sum_j w_{ij} o_{j,t+1} \qquad (3)$$

where w_{ij} is a weight connecting unit i to unit j, and $o_{j,t+1}$ is the output from unit j at time $t+1$. At present, the current model indicates the presence or absence of a letter in the visual field through the activation of a letter unit. The letter unit is connected to its appropriate word unit. The determination of what letter unit is connected to what word units is done *a priori* by the model. The current focus of the model is saccade target selection, and the development of a more complex word recognition module is planned as a further extension to the model.

Saliency Units

The saliency units receive activation from both input and letter units. The input from the letter units represents crosstalk between the "what" and "where" processing pathways, and provides a direct top-down "cognitive" contribution to the evolution of the saliency values for specific regions of the visual field. The saliency unit transfer function is the probability density of the Gaussian distribution with the same parameters as that of the letter units. The units accumulate activity over time and reach a peak of activation after 50 time steps, corresponding roughly to the eye-brain transmission lag (McConkie, 1983; Sereno, Rayner & Posner, 1998).

The role played by the saliency map in the model is to support the target selection process, whereby the word "blob" with the highest activity at a certain point in the processing of a fixation acts as the target for the next saccade (Findlay & Walker, 1999). This can potentially be the currently fixated word, the preceding word, and one or other of the succeeding words. Once the blob with the highest level of activity is selected, a saccade generator module is used to execute a saccade in a way that implements the metrical properties of saccade amplitudes in reading described by McConkie, Kerr, Reddix and Zola (1988; see below).

Word units

Word units receive inputs from their respective letter units, and in turn send activation back to these letter units. There are seven word units, since this is the maximum number of words that were found to be contained in the visual field of width 30 in the text used in the experiments. These units use the following sigmoid transfer function:

$$s(x) = \frac{1}{1 + e^{-\frac{(x-m)}{8}}} \qquad (4)$$

where x is the accumulated net input to the unit and m is, again, 50. This function outputs in the range 0 through 1, with 0.5 for an input of m. The divisor 8 is used to linearise somewhat the S-shaped sigmoid function.

The net input, x, to this equation is slightly more complex than for other units.

$$x_{i,t+1} = x_{i,t} + \frac{\sum_j w_{ij} L_{j,t+1}}{n} + \sum_k w_{ik} W^r{}_{k,t+1} - \sum_m w_{im} W^o{}_{m,t+1} \qquad (5)$$

The terms W^r and W^o denote recurrent inputs and inputs from other words, respectively. Note that the letter input is averaged over word length *n*, so that word length does not affect the rate of activation accumulation, just the average activity of the component letters.[1] The value of the self-recurrent connections is a function of the word's frequency. The higher the frequency the more activation the word receives, and the more rapidly its output peaks. The specific values of the connections are determined by the parameter search mechanism. Note again that the same connection value is used to for each connection type.

The use of word frequency is the only high-level factor that comes into play in this instance of the model. There is scope, however, for using the threshold adjustment of the Fixate Centre unit to implement a strategic modulation of the reading process.

In addition to the self-recurrent connections, the word units are also connected to neighbouring word units with inhibitory connections. This implements a competition for word processing resources. Within this framework, words can be processed in parallel from a given fixation, but one word will tend to dominate at a given point in time. In this way, we have a mechanism for reconciling the different strands of evidence that, on the one hand, appear to suggest a sequential left-to-right processing of words, and on the other, the simultaneous processing of more than one word in a given fixation (see Inhoff & Weger, this volume, for a discussion).

DETERMINING SACCADE METRICS

From the data of McConkie, Kerr, Reddix, and Zola (1988), we have a good idea of the end-point behaviour of the eye in reading. The intended target of a saccade appears to be the centre of the word, and this is attained with varying accuracy as a function of the eye's launch distance from the target.

McConkie *et al.* demonstrated that the distributions of landing sites on a word tended to be Gaussian in shape. The centre of these distributions and their standard deviations appeared to be determined primarily by oculomotor factors. They found a general tendency

[1]This is empirically justified by the aforementioned finding by Inhoff, Radach, Eiter and Juhasz (in press) that parfoveal word length information is not used for lexical processing.

for the eye to land around the centre of the word, and a leftward shift of distribution means with the increase in launch distance. They proposed that the pattern of landing site distributions can possibly be accounted for by five principles: (1) The centre of the word is the functional target of a saccade; (2) a systematic range error causes the eye to be increasingly deviated from this target as a linear function of distance from the launch site; (3) this range error is somewhat less, the longer the eye spends at the launch site[1]; (4) there is a random, Gaussian-shaped distribution of landing sites around the target location; and (5) the spread of this distribution increases as a function of launch distance. These five principles can be summarised in three equations. The first is a linear equation (see Equation 6) describing how the mean landing site (*m*) on a word deviates as a function of launch distance (*d*). Note that both *m* and *d* are defined to be zero at the centre of the targetted word. In the case of a four-letter word, this would be half way between the second and third letter positions.

$$m = 3.3 + 0.49\,d \qquad (6)$$

The second is a cubic equation (see 7) describing the spread of landing positions around *m*.

$$sd = 1.318 + 0.000518\,d^3 \qquad (7)$$

The third is a Gaussian equation (see 8) accounting for the random distribution of landing sites, and for which *m* and *sd* are the parameters.

$$f(x;m,sd) = \frac{1}{\sqrt{2\pi.sd}}e^{-\frac{(x-m)^2}{2.sd^2}} \qquad (8)$$

In the present version of the model, once the activation of the FC unit falls below threshold, these equations are used to determine the amplitude of a saccade aimed at the centre of the word with the highest saliency.

FIXATE CENTRE UNIT

The FC is a single unit with a recurrent connection and connections from the letter units. It implements a Gaussian transfer function identical to that used in the letter and saliency units, so that recurrent inputs and inputs from the letter units gradually cause the unit

[1] This suggestion rested on a trend in the data of McConkie *et al.* (1988) that was not statistically tested. It was not replicated in the data of McConkie, Grimes, Kerr & Zola (1990) and a detailed analysis of a large corpus of individual reading data also did not find evidence in favor of this idea (Radach & Heller, 2000). We therefore decided not to include it in our specification of saccade metrics.

output to rise to a peak and fall away. There is also a stochastic element to the behaviour of the FC unit. While for all other Gaussian equations we use a value of 50 as their mean, the FC unit Gaussian's mean can vary random between 50±10. This effectively speeds up the rise and fall of activity in the FC Gaussian, which permits us to model relatively brief fixations that can occur on a word.

Once the FC output drops below a certain threshold, a saccade is executed to the word location in the saliency map that has the highest saliency value.

Dynamics

The dynamics of the model are typical of the broad class of "interactive activation" models. At the start of a fixation, the 30 element input vector of units is activated with values that are a function of whether there is a character present in a specific location or not, and the eccentricity of that location. As mentioned above, a gamma function is used to weight the inputs.

Once there is a pattern of activity on the input units, the network connections are dynamically configured to ensure that the appropriate letter units connect to the appropriate word units, and vice-versa. Obviously, this is not meant to be analogous to any biological process, and is used here as a computational convenience to reduce the size of the network needed to run the simulation. The default state of the network is for every letter unit to be connected by bi-directional connections to every word unit. The process of configuration that occurs at the beginning of each fixation eliminates spurious connections. Note that the activation values of word and letter units are carried over from one fixation to the next. By this mechanism, spill-over and preview effects are implemented.

With the network configuration complete, the input activation is fed forward to a set of letter units and a set of saliency units, each of which comprises 30 units. There are one-to-one connections from the input units to the letter and saliency units. There are also feedback connections from the word units to the letter units. Because of the use of the Gaussian PDF as the transfer function for the letter and saliency units, the activity of the letter units reaches a peak after a number of cycles of activation. This has been set at 50 cycles, so that the letter unit representing input from the fovea of the visual field will peak after 50 time steps, and will then start to decline. The further one moves away from the fovea, however, the more slowly the level of activation accumulates. The activation of non-foveal letters will reach the same peak value, but will take an increasingly larger number of time steps the further one moves from the fovea.

The letter units receive top-down input from the word units whose level of activity is a function of the *average* letter input from the letter units, and the frequency of occurrence of that word in the language. The more frequent is the word, the more rapidly it is activated, and the more rapidly it asymptotes to an output value of 1.0. Frequency effects are implemented by a positive self-recurrent connection that is proportional to the frequency of the word. Thus the activation levels of high-frequency words rise more rapidly than lower frequency words, but this is also a function of the activity of the letter units, which in turn is a function of the eccentricity of the letters in the visual field. Consequently, visually eccentric high frequency words will be more rapidly identified (i.e., their activity will peak earlier) than their low frequency counterparts.

During the processing of words in a fixation, there is competition between words, mediated by inhibitory connections between word units. Once a word has peaked, it ceases to compete, leaving the way open for other words to complete their processing. In this way, several words can be processed simultaneously, but usually one word takes the lion's share of processing resources.

Activation from the letter units is also sent to the saliency units where it combines with the activation from the input units. Again the transfer function is a Gaussian PDF, which models in one function the accumulation of activation and its decay over time. Areas of high-activation peak and decay more rapidly. Given the varying resolution of the input units (modelled by the gamma function), saliency units receiving foveal inputs will peak and decay more rapidly than other units. So after a certain number of iterations, the saliency values will drop in the foveal regions of the saliency units, implementing a form of "inhibition of return" (Gibson & Egeth, 1994).

Activity from the letter units is passed to the fixation centre unit. This unit acts as a spatially undiscriminating summary of saliency unit activity. So it will, in principle, trigger a refixation, regressive or progressive saccade, depending on where the saliency maximum is located.

Parameter fitting

In order to determine an appropriate set of weights for the various unit connections in the simulation, a parameter search algorithm was devised. The algorithm used was based on the Alopex neural network-learning algorithm (Unnikrishnan & Venugopal, 1994). Instead of using an *error gradient* to guide the changes in weight (or parameter) values, Alopex uses *local correlations* between changes in individual weights and changes in a global error measure. The algorithm does not make any assumptions about the transfer functions of individual units, and does not explicitly depend on the functional form of the error measure. This makes it ideal for parameter fitting, where we might want to combine a number of factors into a complex cost function that we wish to minimize. Thus, we can ensure the algorithm selects a set of parameters that satisfies the global temporal and spatial characteristics of reading, specifically fixation duration and saccade length.

The algorithm is initially *stochastic* in its search, and uses a "temperature" parameter in a manner similar to that in simulated annealing (Kirkpatrick, Gelatt, & Vecchi, 1983) to gradually make the search more deterministic as the algorithm converges on a desirable set of parameters[1].

The parameter search involves making small perturbations (e.g., ±0.01) to the parameters based on whether the previous change resulted in a reduction of the cost function. The parameter $w_{ij}(n)$ in Equation (9) below, refers to the connection between units i and j at time n. This parameter is perturbed by $\pm\delta$, where δ is a constant value.

$$w_{ij}(n) = w_{ij}(n-1) + \delta_{ij}(n) \tag{9}$$

$$\delta_{ij}(n) = \begin{cases} -\delta \text{ with probability } p_{ij}(n) \\ +\delta \text{ with probability } 1\text{-}p_{ij}(n) \end{cases} \tag{10}$$

However the changes depend probabilistically on the cost function value, as can be seen from Equation (11).

[1] The temperature term in simulated annealing has the effect of increasing the stochasticity of the parameter search when the error is high. It's analogous to the process used in steel manufacture of heating and slow cooling of the metal to encourage the formation of more stable crystalline structure and thus increase the metal's strength.

$$p_{ij}(n)=\frac{1}{1+e^{\delta\frac{\Delta E(n)}{T(n)}}} \tag{11}$$

This probability is modulated by a temperature variable (Equation 12), which is derived from the overall cost function value. As this value decreases, the selection of parameter change moves from being stochastic to deterministic.

$$\begin{aligned} T(n) &= \frac{\delta}{N}\sum_{n'=n-N}^{n-1}\left|\Delta E(n')\right| && \text{if n is a multiple of N,} \\ T(n) &= T(n-1) && \text{otherwise} \end{aligned} \tag{12}$$

where δ is the constant parameter change (usually around 0.01), and *ΔE = E(n-1)-E(n-2)* is the change in error between the two previous iterations.

The Alopex algorithm shows reasonable convergence, but is not as efficient as a gradient descent based algorithm. Nonetheless, the flexibility it provides in the specification of arbitrarily complex cost functions is worth the slower convergence. This is especially true when the number of parameters to be estimated is relatively small.

In the case of the model described in this chapter, there were nine free parameters to be estimated, furthermore, the sign of seven of these was constrained to be positive, and the maximum absolute value permitted was 10.0. The parameters were the following connections: input-to-letter, input-to-saliency, letter-to-word, letter-to-saliency, word-to-letter, word-to-word, word-to-fixate centre, fixate centre -to- fixate centre. All except word-to- fixate centre were constrained to be positive in sign.

The text used for the parameter fitting was a 2,500 word German text on the topic of the Inuit. The cost function used was:

$$Cost=(n-1)^2+\left(\frac{1}{m}-1\right)^2+((p-11)-8)^2+(t-0.25)^2 \tag{13}$$

where *n* is the number of peaks (at a minimum for just one peak), *m* is the maximum value of the peak (at a minimum when 1.0), *p* is the location of the maximum peak (at a minimum for 8 characters to the right of the fovea), and *t* is the threshold for the FC unit (the desired threshold is 0.25).

<Insert Figure 3 about here>

Evaluation

The model can deal with a variety of low-level reading phenomena in an integrated and parsimonious manner. In this section, we will demonstrate the operating principles of the model and how they can account for such phenomena as preview effects, spillover effects, refixations, regressions, and word skipping. The advantage of this model over others is, we believe, the variety of phenomena that can be accounted for with letter level accuracy within a rather simple computational framework.

Figure 3 illustrates how the word units, saliency units, and fixate centre interact on a given fixation. Once the activity in the fixate centre falls below a certain threshold, a saccade is triggered to the word with the highest saliency value, irrespective of where it is. In this case, the word with the highest saliency is $n+1$, resulting in a progressive interword saccade. Also illustrated in Figure 3 is the mechanism mediating preview effects. These effects arise through the activation value of the new word with the rising level of activity is being carried over to the next fixation. Note that preview does not result from the disengagement and re-engagement of an attentional mechanism. Rather we propose a continuous mechanism that dynamically modulates the processing load across the words in the fovea and immediate neighbourhood as a function of their relative difficulty. Thus, more than one word can be processed at a given time, though there is competition for lexical processing resources between words. Note that once a word reaches its asymptotic value, it no longer competes with other words.

<Insert Figure 4 & 5 about here>

The model is capable of accounting for the modulation of preview benefit by the difficulty (in this case, low frequency) of the currently fixated word (see Figure 4). By modeling the processing of the words as a continuous asymptotic process, we can account for the dynamic interplay between the processing of word n and word $n+1$. In Figure 4, the frequency of word n is varied. Where word n is a high frequency word, the level of word activity rapidly rises to a peak, thus removing it from competition with word $n+1$, and allowing it, in turn, to be processed. When word n is a low frequency word, less progress is made in processing word $n+1$, thus reducing any preview benefit for it.

Spillover effects from the processing of the previous word on the currently fixated word can be accounted for in precisely the same way. In Figure 5, we see that the processing of the low-frequency word has not asymptoted prior to the fixation. Recall that the trigger for executing a saccade is a drop in the level of activity of the fixate centre below a certain threshold. It bears no *direct* relationship to the successful, or otherwise, processing of the currently fixated word. If a high frequency word precedes the current fixation, there is less likely to be processing of that word continuing into the next fixation.

<Insert Figure 6 about here>

Refixations

In Figure 6, we show the performance of the time course of activation over the letter field when an eight-letter word is fixated for the first time at its last letter. Note that there has been no preceding fixation of this word or the one preceding it. The situation is equivalent to there having been a long saccade from the right to this point in the text. There is a build up of activation in two competing word targets on the saliency map, with the currently fixated word (*beweisen*) marginally winning the competition. This word is then selected as the target for the next saccade.

<Insert Figure 7 about here>

Regressions

In figure 7, we have a similar graph showing the build up of activation taking place over the location of the preceding word. Again, there was no fixation on the currently fixated or preceding word prior to this one. The fixation position is at the beginning of an eight letter high-frequency word. Consequently it drops in saliency fairly rapidly, leaving the preceding word the most salient target. If a saccade is triggered at this point, the result would be a regression to the preceding word.

<Insert Figure 8 about here>

Likelihood of fixating a word ("word skipping")

Figure 8a and 8b illustrate how the model can account for the likelihood of fixating a word as a function of its frequency. As mentioned above, we do not believe that there is any default tendency to aim a saccade at each word on a line of text; therefore we put the traditional term "word skipping" in quotation marks. Note that in this simulation example, each of the graphs shows the second of two fixations on the word or within the vector, the preceding fixation having been on the word to left of the current fixation (fixation locations are indicated by arrows). In Figure 8a, we can see that the high frequency word in the right parafovea does not receive a fixation, whereas in 8b the low frequency word gets fixated. This effect is achieved by the slower accumulation of activation for the low frequency word, and the subsequent maintenance of activity in the saliency map in the region of the low frequency word.

Conclusion

The Glenmore model has been described in outline, and some qualitative and quantitative simulation results have been presented. We have proposed a radical departure from the more traditional attention shift models of eye movement control reading. Our model does not use a notion of attention shift, but rather permits the processing of words in parallel, with a limited amount of competition between words in a given fixation. The model decouples the decision about when to move the eyes from the word recognition process, but allows for a substantial influence of linguistic processing on the movement decision. More specifically, the time course of activity in a "fixate centre" determines the triggering of a saccade in a way that is co-determined by a random component and ongoing processing on the letter and word level. The other main feature of the model is the use of a saliency map that acts as an arena for the interplay of bottom-up visual features of the text, and top-down lexical features. These factors combine to create a pattern of activation that selects one word as the saccade target.

As we have emphasized in the introductory section, our theoretical approach is closely related to the more general theoretical framework developed by Findlay & Walker (1999). While our model is in principal quite similar to their conception, it is much more precise with respect to a number of reading-specific mechanisms. It is less detailed in other respects, and in some ways it departs from specific suggestions made by these authors. The most important difference lies in the mechanisms and dynamics of saccade triggering. In both models it is assumed that a saccade is triggered when activity in the fixate centre falls below threshold. However, according to Findlay & Walker, increased activity in the move centre (the saliency map) promotes a decline of activity in the fixate centre via reciprocal inhibitory connections. In addition, a decrease in activity in the fixate centre can also be triggered "via descending influences" from higher processing centres (p. 663). However, their ideas on the nature of these higher order influences remain vague, and need to be specified in a model on a task as specific as reading.

Why is there no reciprocal connection between the saliency map and the fixate centre in the Glenmore model? Findlay and Walker's theory is concerned with explaining the latency of a saccade that is generated in a way that can account for a number of experimental effects like the gap effect or the remote distractor effect that are all variations of a dramatically changing visual input. In the context of explaining the latency of a single saccade under such conditions, competition between potential targets on the visual level plays a key role. There is a lively discussion on how to model these influences up to the possibility that such saccades may be elicited when a threshold level of activation is surpassed in a spatially selective move system rather then a fixate centre (Dorris & Munoz, 1999).

In reading, the situation is quite different in that the visual environment is stable and consists of well defined elements in a highly structured spatial arrangement. Hence, there is no active competition of targets on the level of visual processing other then what is captured with assigning initial saliency values to the elements in the perceptual span vector. The rest of the saliency dynamics comes entirely from processing on the letter and word level. It would therefore not add to the explanatory power of the model to include an extra connection from the saliency vector to the fixate centre. The optimal way to model the substantial influence of linguistic processing on fixation duration in reading appears to be via the implemented transfer of integrated activity from the letter processing vector to the fixate centre. This mechanism has the advantage of accounting in a parsimonious way for effects on *both* the orthographic and lexical processing level on the duration of fixations.

Within the modelling framework described in this chapter, it is possible to account within one mechanism for preview and spillover effects, and for regressions, progressions, and refixations. Even with this simple model, a range of phenomena that have hitherto proved problematic to account for, have been accommodated in a way that appears to be in harmony with neorophysiological constraints to the degree that such constraints can be made explicit for the task of reading. How does the Glenmore model compare to the similar approaches that have recently been put forward by Yang & McConkie (2001) and by Engbert, Longtien & Kliegl (2002; see also the chapters by McConkie & Yang, and by Kliegl & Engbert in this volume).

The interaction/competition theory by Yang & McConkie (2001) is also based the theoretical framework suggested by Findlay and Walker (1999). The theory is very explicit about the specific mechanisms of saccade triggering and the resulting distributions of fixation durations which they primarily attribute to non-cognitive factors. In comparison the Glenmore model, the I/C theory is less specific about the spatiotemporal dynamics of saliency within the perceptual span. It also does not attempt to model littler level and word level processing in any explicit way. On the other hand, it makes relatively explicit claims about ways in which cognition can influence saccade initiation times. There are three possibilities for such influences, referred to as processing-based inhibition, parametric adjustment and late-acting cognitive control. The first two of these mechanisms could be accounted for within the Glenmore model. Yang & McConkie (2001) presented text with occasional non-text letter strings or pseudowords and observed that saccades normally triggered after fixation durations of more than about 200 ms were substantially delayed. They propose "that these non-text letter strings or pseudowords create processing problems at some level of text processing, and that the brain centers being disturbed send a distress signal that results in inhibition in the saccadic system (McConkie & Yang, this volume)". In the Glenmore model, entering a pseudowords into the letter processing vector would make this string an attractive target for fixation (via transfer of processing activation to the saliency map), with associated long-lasting activation of the fixate center, resulting in delayed saccade onsets. The second mechanism for cognitive influences proposed in the I/C theory, parametric adjustment, is similar to accounting for strategic influences on reading behavior that we have briefly discussed earlier in this chapter. It can be accounted for in the Glenmore model by adjusting the critical threshold in the fixate centre. In sum, we see the Interaction/Competition theory in a complementary relation to our own approach, the major difference being our less extreme viewpoint concerning the relative importance of visual vs. cognitive factors.

As mentioned in the introductory section, the Glenmore model has striking similarities with the SWIFT-model proposed by Engbert, Longtien & Kliegl 2002 (see Kliegl & Engbert, this volume for the latest version). Kliegl & Engbert emphasize that their model is based on three principles: a partial separation of saccade timing from saccade target selection, spatially distributed lexical processing, and autonomous saccade generation with inhibition by foveal word processing. One major difference is that in Glenmore linguistic processing is implemented at the letter and word level and that the influence of processing on the timing of saccade triggering is in terms of the integrated processing activity within the perceptual span rather than exclusive processing of the foveal word. Moreover, in SWIFT the decision which word should be the target of the next fixation is stochastically determined as a function of lexical activity, with the word having the largest current lexical activity being the most likely target. In contrast, the way of modeling a hierarchy of potential saccade targets in Glenmore is via combined visual information and processing dynamics in a spatial saliency map. The SWIFT model is much more specific with respect to the time line of saccade programming, considering in detail various possibilities of temporal overlap in the processing of successive saccades. This is an issue that needs to be reconsidered for Glenmore, including the possibility that our assumptions regarding a simple relation between saccade triggering and execution were too naïve.

Seen in conjunction, theories and models like I/C, SWIFT and Glenmore represent a class of approaches that may differ greatly in detail but appear to share a common philosophy. However, as documented in the chapter by Pollatsek, Reichle & Rayner in this volume, there are also good arguments in favour of their view that reading saccades are triggered on the basis of sequential lexical processing. This range of opinion is likely to generate lively theoretical discussions and to provoke empirical studies to test predictions associated with various models. This may contribute to replacing the low level vs. high level controversies that have dominated the last decade with a more productive debate on *how* visuomotor and cognitive elements combine to determine the movements of our eyes while we read.

References

Brysbaert, M. & Vitu, F. (1998) Word skipping: implications for theories of eye movement control in reading. In G. Underwood (Ed.) *Eye Guidance in reading and scene perception*. Elsevier, Amsterdam, pp. 125-148.

Carpenter, R.H.S. (2000). The neural control of looking. *Current Biology, 10,* R291-R293.

CELEX German Database. Release D25. Computer software. Nijmegen: Centre for Lexical Information, 1995.

Deubel, H., O´Regan, K. & Radach, R. (2000). Attention, information processing and eye movement control. In Kennedy, A., Radach, R., Heller, D. & Pynte, J. (Eds). *Reading as a Perceptual Process.* Oxford: Elsevier.

Dorris, M. C. & Munoz, D. P. (1999). The underrated role of the "move system" in determining saccade latency. *Behavioral and Brain Sciences 22 (4)*, 681-682.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621-636.

Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics, 85,* 77-87.

Findlay, J. M. & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral & Brain Sciences 22 (4),* 661-721.

Gibson, B.S. & Egeth, H.E. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics, 55 (3)*, 323-339.

Grainger, J. (2000). From Print to Meaning via Words? In Kennedy, A., Radach, R., Heller, D. & Pynte, J. (Eds). *Reading as a Perceptual Process.* Oxford: Elsevier.

Grainger, J. & Jacobs, A.M. (1998). On localist connectionism and psychological science. In J. Grainger & A.M. Jacobs (Eds). *Localist Connectionistic Approaches to Human Cognition.* Mahaw, NJ: Earlbaum, pp. 1-38.

Inhoff, A.W., Lemmer, S., Eiter, B. & Radach, R. (2001). Saccade programming and parafoveal information use in reading: More evidence against attention-shift models. Poster, *11th European Conference on Eye Movements*, Turku, Finland.

Inhoff, A.W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming in reading. In A. Kennedy, R. Radach, D. Heller and J. Pynte (Eds.), *Reading as a Visual Process.* Oxford: Elsevier.

Inhoff, A., Radach, R., Eiter, B. & Juhasz, B. (in press). Parafoveal Processing: Distinct Subsystems for Spatial and Linguistic Information. *Quarterly Journal of Experimental Psychology*.

Jacobs, A. M. (2000). Five questions about cognitive models and some answers from three models of reading. In: In A. Kennedy, R. Radach, D. Heller and J. Pynte (Eds.), *Reading as a Perceptual Process.* Oxford: Elsevier.

Just, M.A. & Carpenter, P.A.(1987). *The psychology of reading and language* Blanchard, H.E. (1985). A comparison of some processing time measures based on eye movements. *Acta Psychologica, 58,* 1-15.

Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and beahvior.* Cambridge, MA: MIT Press.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing tradeoff. In G. Underwood (Ed.), *Eye guidance in reading and scene perception*. Oxford: Elsevier, pp. 149-180.

Kennedy, A., Radach, R., Heller, D. & Pynte, J. (Eds., 2000). *Reading as a Perceptual Process.* Oxford, Elsevier.

Kerr, P. W. (1992). Eye movement control during reading: The selection of where to send the eyes. *Unpublished doctoral dissertation*. University of Illnois at Urbana-Champaign.

Kirkpatrick, S., Gelatt Jr., C. D., Vecchi, M. P. (1983). *Optimization by simulated annealing,* Science, *220*, 671-680.

LaBerge, D. & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96*, 101-124.

McClelland J.L., & Rumelhart D.E. (1981): An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. *Psychological Review, 88,* 375-407.

McConkie, G. W., Kerr, P. W. & Dyre, B. P. (1994). What are 'normal' eye movements during reading: Toward a mathematical description. In: Ygge J, Lennerstrand G (eds), *Eye movements in reading.* Pergamon, Oxford, pp 315-328.

McConkie, G. W., Underwood, N. R. , Zola, D. & Wolverton, G. S. (1985). Some temporal characteristics of processing during reading. *Journal of Experimental Psychology: Human Perception and Performance 11*: 168-186

McConkie, G. W. & Zola, D. (1987). Visual attention during eye fixations while reading. In M. Coltheart (Ed.), *Attention and Performance* (Vol. 12, pp. 385-401), London, NJ: Erlbaum.

McConkie, G. W. (1983). Eye movements and perception during reading. In K. Rayner (Ed.), *Eye movements in reading. Perceptual and language processes.* Academic Press: New York, pp. 65-96.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578-586.

McConkie, G. W., Grimes, J. M., Kerr, P. W. & Zola, D. (1990). Children's eye movements during reading. In J. F. Stein (Ed.), *Vision and Visual Dyslexia*. MacMillan.

McConkie, G. W., Kerr, P. W., Reddix, M. D., and Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28,* 1107-1118.

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception and Psychophysics, 46,* 245-253.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance* 10: 667-682.

Munoz, D. P. (in press). Organisation of cortical and subcortical brain areas contributing to the control of visual fixation and saccade initiation. In Hyönä, J., Munoz, D., Heide, W. & Radach, R. (Eds.), *The Brain`s Eyes: Neurobiological and Clinical Aspects of Oculomotor Research.* Progress in brain research. Oxford: Elsevier Science.

O'Regan, J. K. (1990). Eye movements in reading. In E. Kowler (Ed.) *Eye movements and their role in visual and cognitive processes*. Elsevier: Amsterdam, pp. 395-453.

O'Regan, J. K., Vitu, F., Radach, R. and Kerr, P. (1994). Effects of local processing and oculomotor factors in eye movement guidance in reading. In J.Ygge and G. Lennerstrand (Eds.), *Eye Movements in Reading.* New York: Pergamon, pp. 329-348.

Port, R.F., & Van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition.* Cambridge, MA: MIT Press.

Radach, R. & Heller, D. (2000). Relations between Spatial and Temporal Aspects of Eye Movement Control. In Kennedy, A., Radach, R., Heller, D. & Pynte, J. (Eds). Reading as a Perceptual Process. Oxford, Elsevier.

Radach, R., Inhoff, A. W. & Heller, D. (2002). The role of attention and spatial selection in fluent reading. In Witruk, E. Friederici, A. D. & Lachmann, T. (Eds.). *Basic function of language, reading, and reading disability.* Boston: Kluwer, pp. 137-154.

Radach, R. (1999). Top-down influences on saccade generation in cognitive tasks. *Behavioral and brain sciences,* 22.4, 697-698.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading.* Englewood, NJ: Prentice-Hall.

Reichle, E. D., Pollatsek, A., Fisher, D. L. & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105* (1), 125-157.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research, 39,* 4403-4411.

Reichle, E., Rayner, K & Pollatsek, A. (in press). Comparing the E-Z reader model to other models of eye movement control in reading. Behavioral and Brain Sciences.

Reilly, R., (1993a). A connectionist attentional shift model of eye movement control in reading. In: Proceedings of the fifteenth annual conference of the Cognitive Science Society, University of Colorado at Boulder, pp. 860-865.

Reilly, R. (1993b). A connectionist framework for modeling eye-movement control in reading. In G. d'Ydewalle & J. Van Rensbergen (Eds.), Perception and cognition: Advances in eye-movement research. Amsterdam: Elsevier.

Reilly, R.G., & O'Regan, J.K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research,* 38, 303-317.

Rumelhart, D.E. & McClelland, J.L. (1982). An Interactive Activation Model of Context Effects in Letter Perception: Part 2. The Contextual Enhancement Effect and Some Tests and Extensions of the Model,. *Psychological Review, 89*, 60-94.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a timeline of processing during reading: Evidence from eye movements and event-related potentials. *NeuroReport,* 9, 2195-2200.

Unnikrishnan, K.P., & Venugopal, K.P. (1994). Alopex: A correlation-based learning algorithm for feedforward and recurrent neural networks. *Neural Computation*, *6*, 469-490.

Wurtz, R. H. (1996). Vision of the control of movement. *Investigative Ophthalmology & Visual Science*, 37 (11), 2131-2145.

Yang, S.-N., & McConkie, G.W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research.* 41, 3567-3585.

Figure Captions

Figure 1: Model overview

This figure represents the main components of the Glenmore model. The circles represent connectionist components, the rectangle a non-connectionist component. Connections with filled circles represent negative connections, those with arrows positive ones. Note that the negative connection from words to letters act to maintain activity in the letter units, when those units have a cumulative Gaussian transfer function. This is because the negative top-down values will impede the rate at which the activation values saturate as a function of their bottom-up inputs.

Figure 2: Model detail

This figure is a schematic representation of the Glenmore's internal connectivity. The transfer functions of the relevant units are represented graphically in boxes adjacent to them. Activity in the 30 input units, scaled by the gamma function to represent variability in spatial resolution, is propagated to the saliency map (the "where" pathway) and to the letter units (the "what" pathway). Activity in the letter units feeds forward to the word units, which in turn feed activation back to the letter units. Note that the feedback from words to letters is negatively weighted, so that letters receiving a large amount feedback have there activation maintained for longer. The recurrent connections on the word units are used to implement word frequency effects. The more familiar or frequent the word, the more rapidly its activation will rise and decay. The fixate centre unit takes input from the letter units. When that activity falls below a certain threshold, a saccade is triggered to the word with the largest peak in the saliency map. There is also a stochastic component to this process that will cause the activity of the fixate centre to decay at varying times.

Figure 3: Time course of processing

This figure represents the time course of processing in a number of the components of the network (the letter units are not shown to reduce the complexity of the figure). The top panel represents the time course of activation of two word units. The activation of word units is carried over from fixation to fixation. The vertical line down through the panels indicates the time at which a saccade was triggered. At that time, the word represented by the green line has asymptoted and the activation value of the second word has started to rise. This level of activation represents the preview benefit for that word when it comes to the start of the next fixation. The second panel is a spatio-temporal representation of the activity levels across the saliency map. The bars indicate regions in saliency map that have activation values greater than a 0.5 at a given time step. When a saccade is triggered, the word with the highest saliency peak is the saccade target. The bottom panel is a representation of the time course of activity of the fixate centre. When this falls below an adjustable threshold, a saccade is triggered. The activity of the fixate centre is a function of the activity of the letter units.

Figure 4: Preview modulation by frequency of fixated word

This figure represents schematically the interaction between word frequency and peripheral preview. The time course of word activation for two pairs of words is represented, where word *n* in each pair (green curve) is either a low frequency word (top panel) or a high-frequency word (bottom panel). Because there is competition between active word units (until they reach asymptote), if word *n* remains active longer it will limit the processing of word *n*+1 (blue curve). This is represented by differences in the rise time of the word *n+1*, as a function of the frequency of word *n*.

Figure 5: Spillover effects from preceding word

This figure schematically represents the production of spillover effects. The time course of word activation for two pairs of words is represented, where word *n* in each pair (green curve) is either a low frequency word (top panel) or a high-frequency word (bottom panel). As illustrated in the top panel, if a saccade occurs before word *n* reaches asymptote, its activation will carry over to the next fixation, competing for processing in the succeeding fixation. This will not happen where word *n* is a high frequency word (bottom panel).

Figure 6: Modelling a refixation

This figure shows two time-slices of activation in the saliency map for a fixation on the last letter of the word “Beweisen.” Note that neither the current nor previous word had previously been fixated. The activation peak over the current word is marginally ahead of word *n+1*. A saccade triggered at either 200 or 280 simulated msecs would, therefore, result in a re-fixation of the current word.

Figure 7: Modeling a regressive fixation

This figure shows two time-slices of activation in the saliency map for a fixation on the first letter of the word “zwischen.” As with Figure 6, neither the current nor previous word had previously been fixated. The activation peak over the preceding word results in a regressive fixation to word *n-1*.

Figure 8: Frequency effects on word fixation likelihood

This figure shows the effect of the frequency of word *n+1* on the likelihood of it being fixated. The black arrow indicates the current fixation. Note that for both figures the preceding fixation was on the last letter of the preceding word, indicated by the red arrow. Because “nicht” is a higher frequency word than “Namen”, and because saliency is influenced by word frequency, the saliency location of the more frequent word declines more rapidly than the less frequent word. A saccade triggered at 240 simulated msecs to the word location with the highest peak will result in a higher frequency *n+1* not being targeted for a fixation (8a), but with the opposite the case for a lower frequency *n+1* (8b).

Figure 1: Model overview

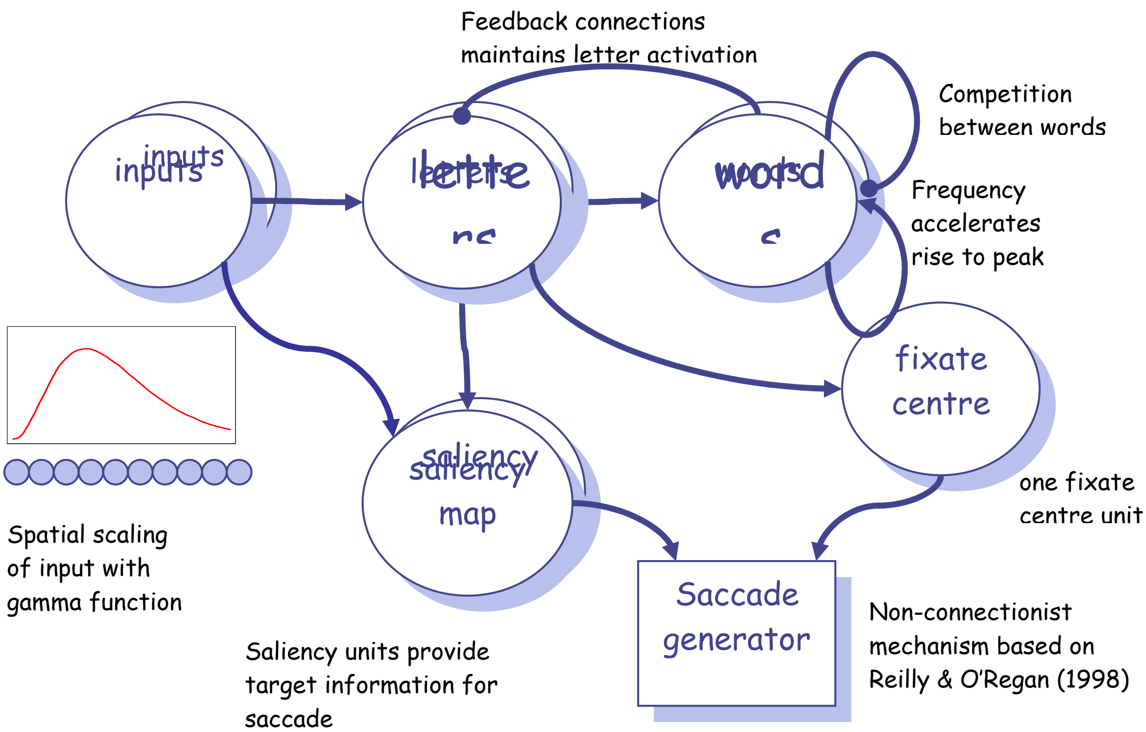

Figure 2: Model detail

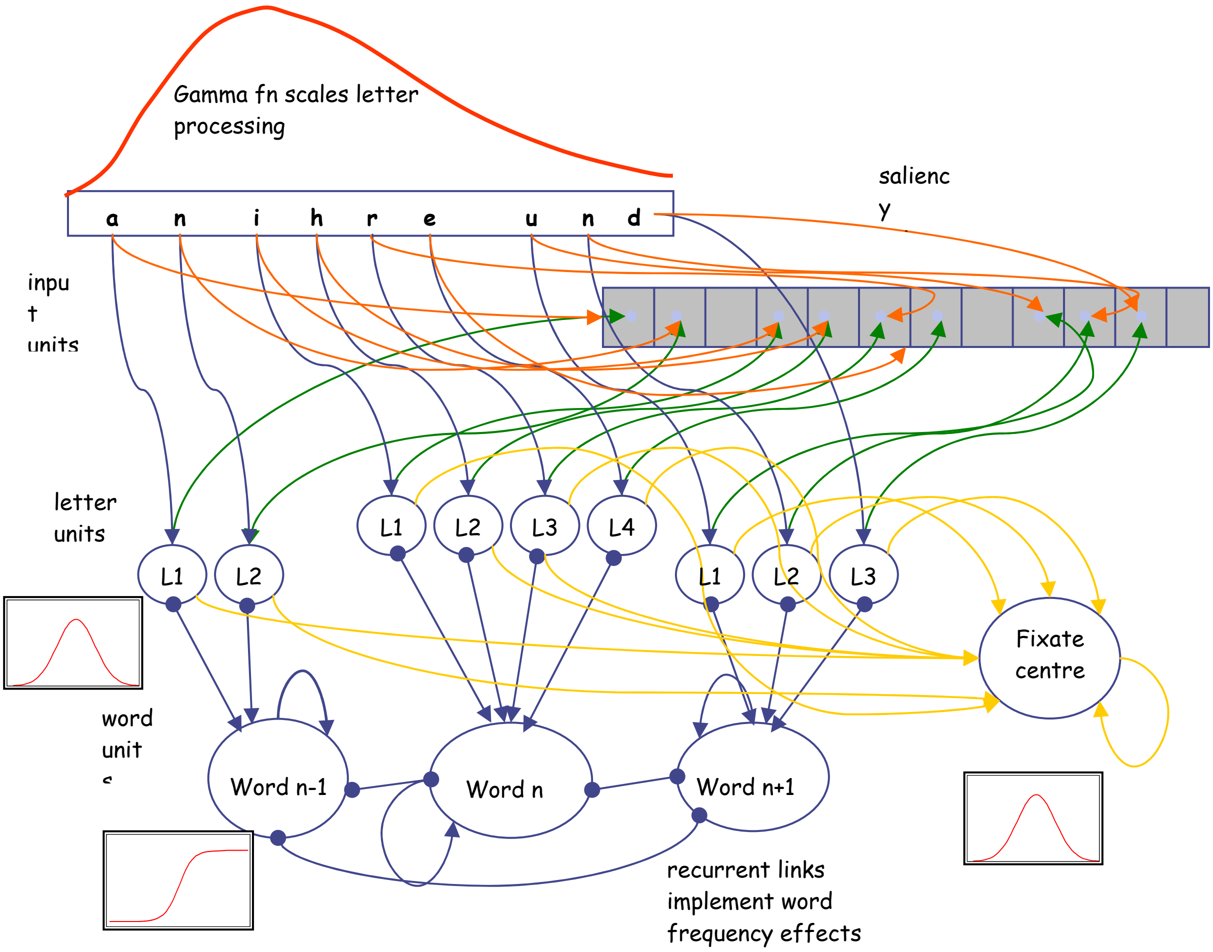

Figure 3: Time course of processing

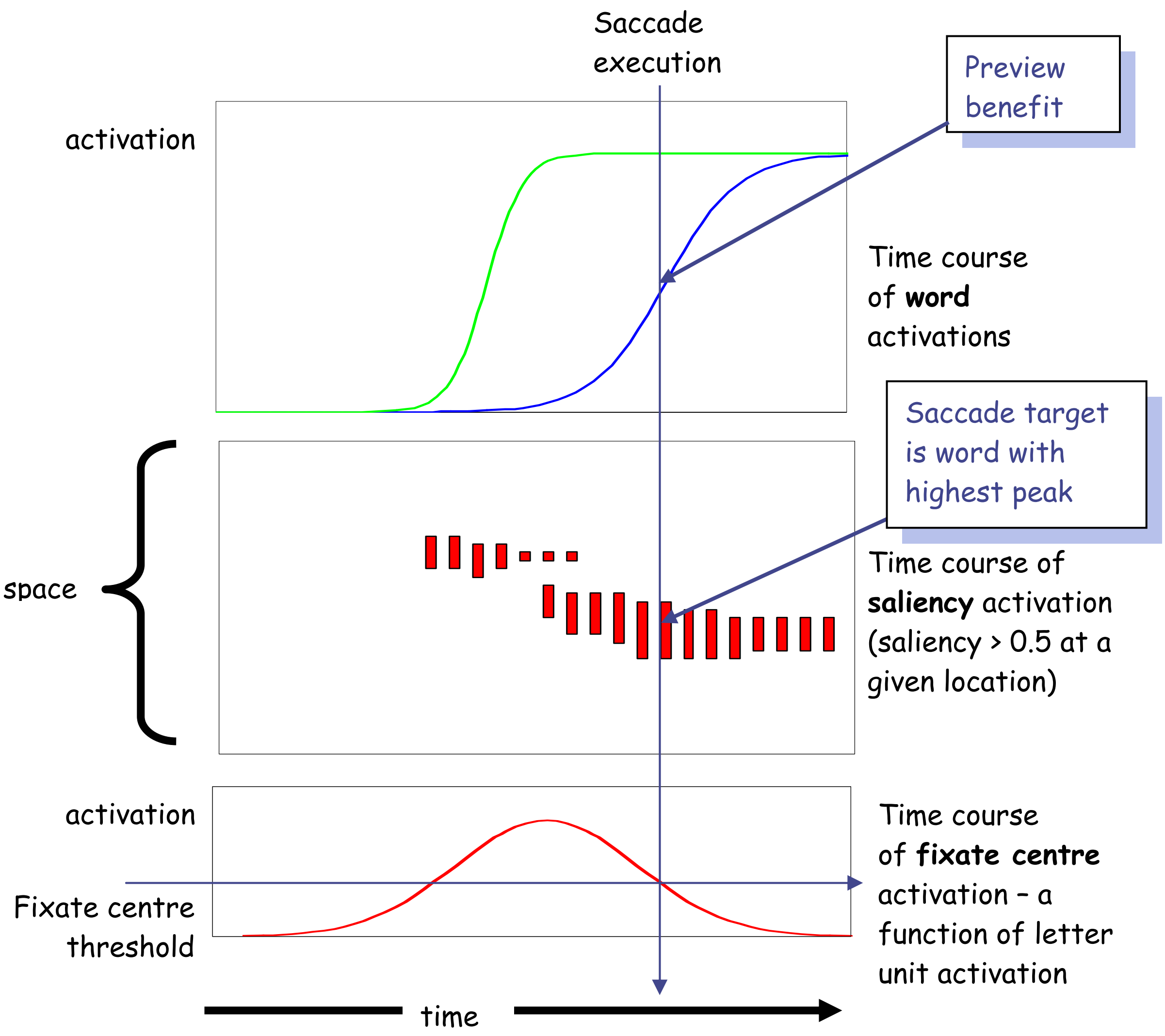

Figure 4: Preview modulation by frequency of fixated word

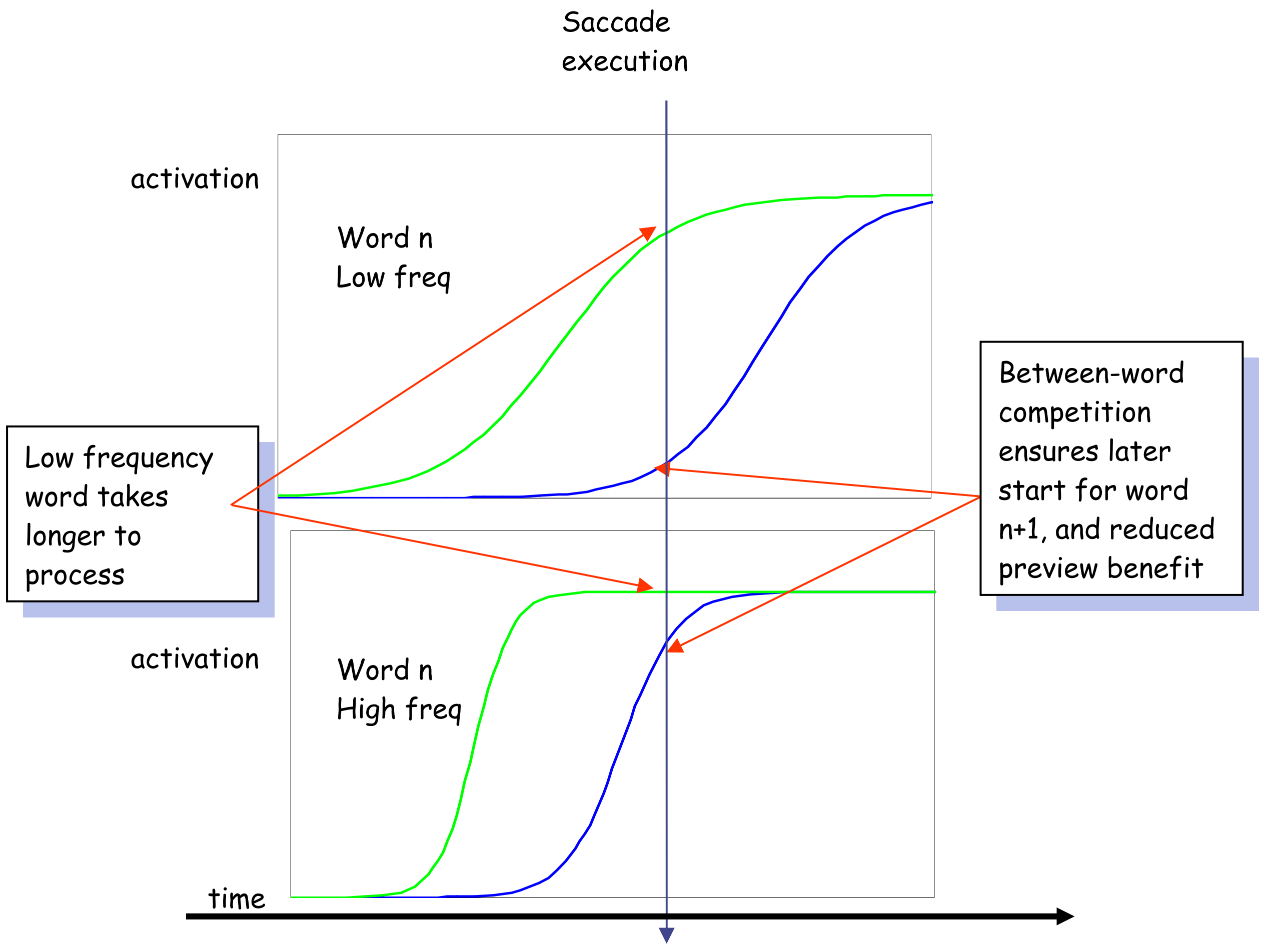

Figure 5: Spillover effects from preceding word

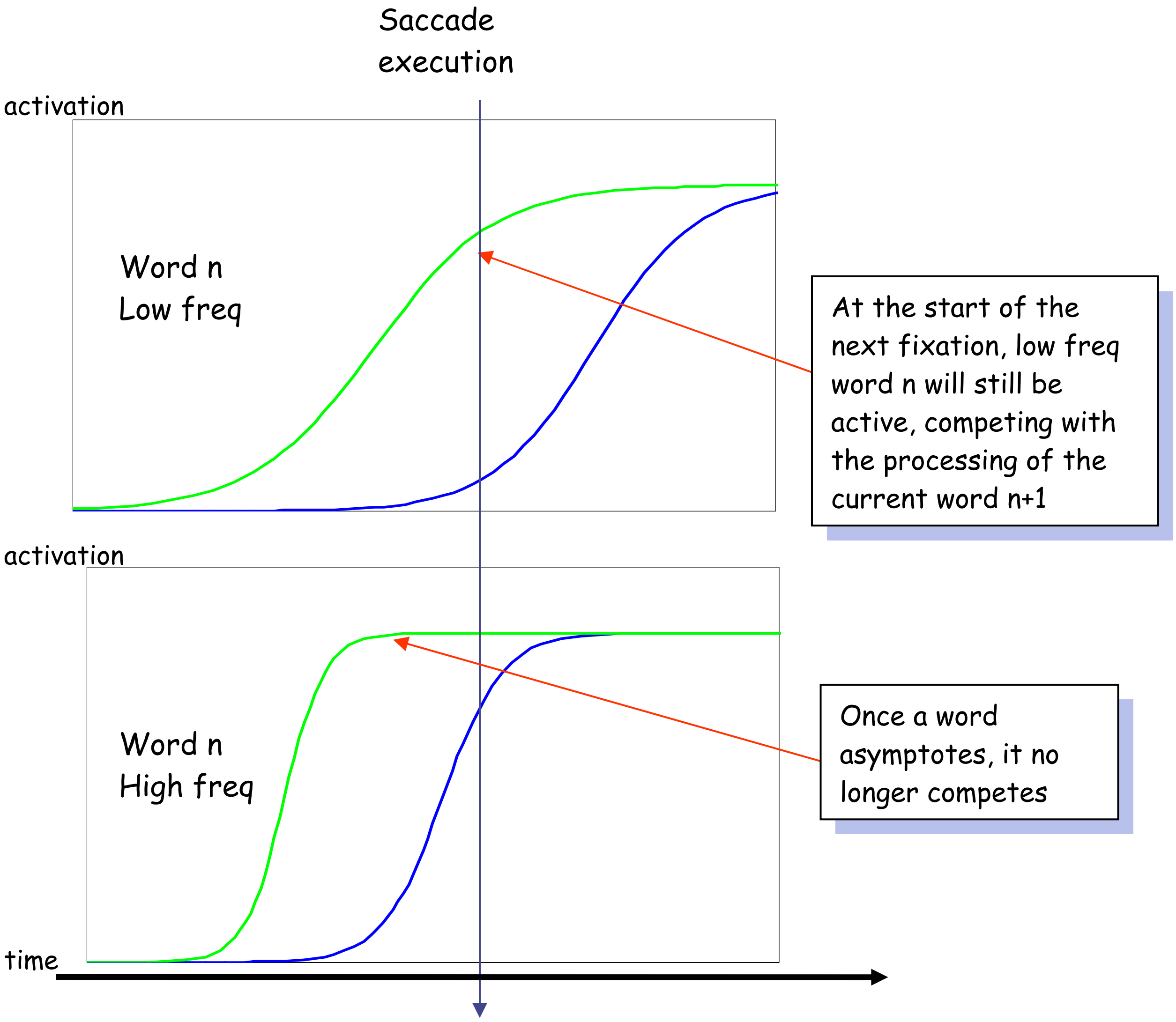

Figure 6: Modelling a refixation

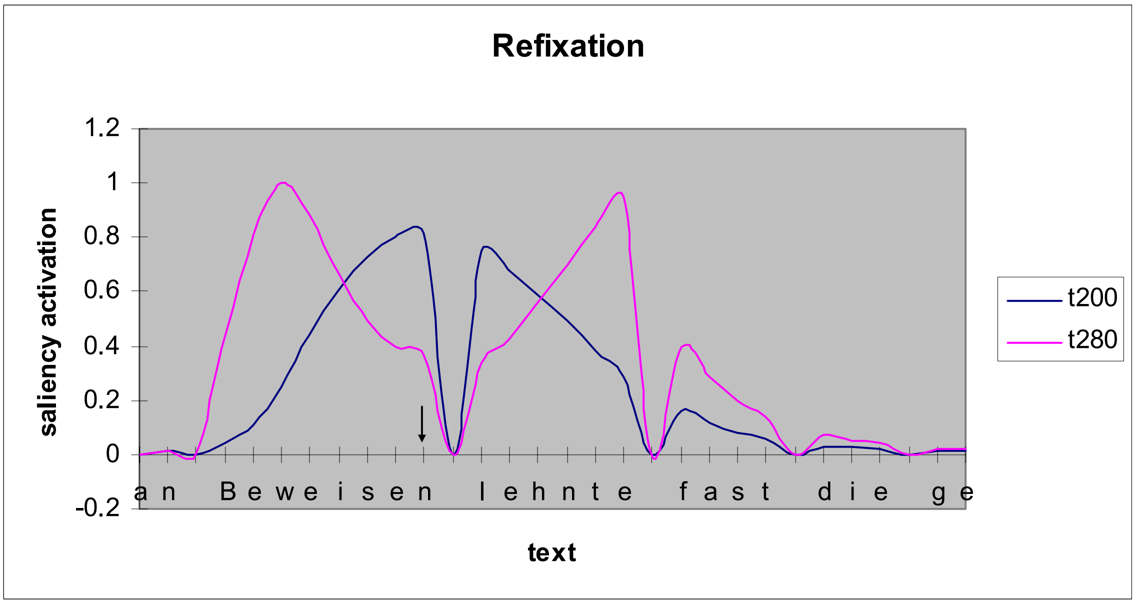

Figure 7: Modeling a regressive fixation

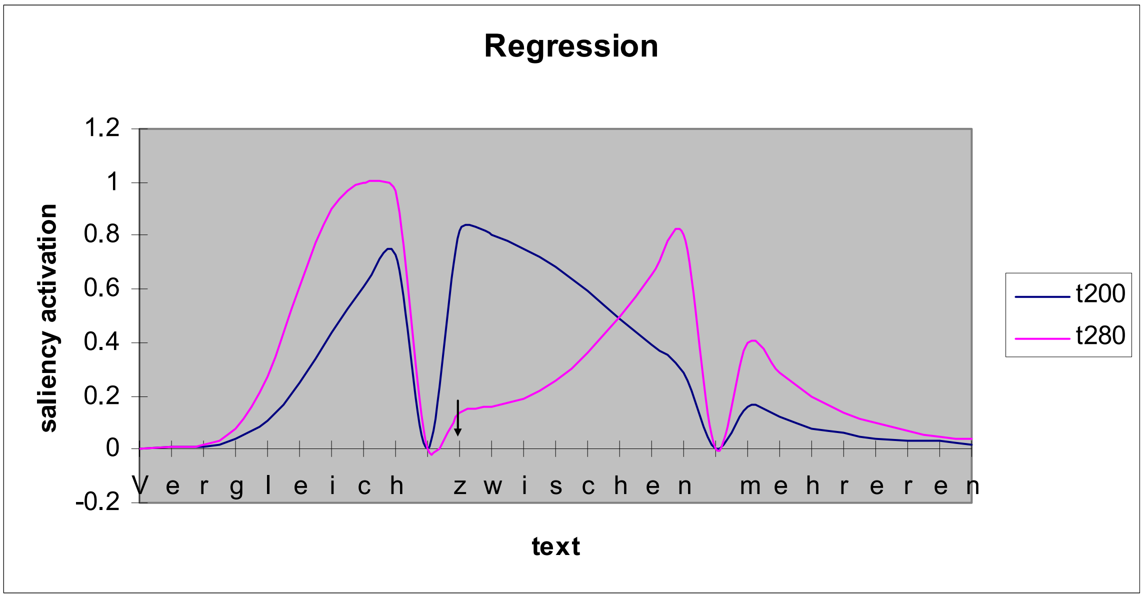

Figure 8(a): Word *n+1* is a high frequency word

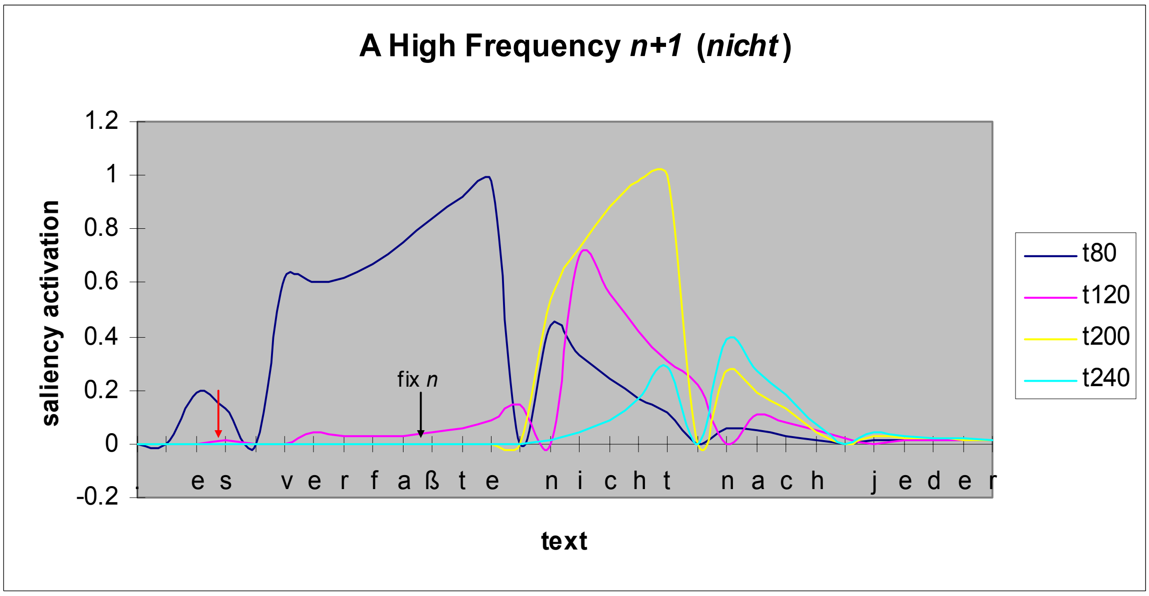

Figure 8(b): Word *n+1* is a low frequency word